Classification of Cataract Fundus Image Based on Deep Learning

Yanyan Dong[1], Qinyan Zhang[1],Zhiqiang Qiao[1],Ji-Jiang Yang[2]

1. Automation School, Beijing University of Post and Telecommunications, Beijing, 100876
2. Research Institute of Information Technology, Tsinghua University, Beijing, 100084

Corresponding author: yangjijiang@tsinghua.edu.cn

Abstract*—Cataract is a dulling or clouding of the lens inside the eye. Which is one of the most common diseases that might cause blindness. Considering the damage impact of cataract, we propose to use computer science for automatic cataract detection, which is based on the classification of retinal image. This method focuses on the feature extraction step of retinal image. Firstly, the maximum entropy method is used to preprocess the fundus images. Next, we use deep learning network which is based on Caffe to automatically extract more distinctive features of fundus images. Last, several representative classification algorithms are used to identify automatically extracted features. Comparing to features extracted by deep learning and wavelet feature extracted from retinal vascular, SVM(support vector machines) and Softmax are used for cataract classification. Finally, cataract images are classified into normal, slight, medium or severe four-class. Through comparing with the results of classification, the feature extracted from deep learning which is classified by Softmax get better accuracy. The results demonstrate that our research on deep learning is effective and has practical value.

***Keywords* —Cataract; Deep learning; Caffe ; Retinal vascular feature ;Feature extraction; Softmax**

I. INTRODUCTION

With the popularity of electronic devices in modern society, people have growing demand on using eyes. So eye protection becomes increasingly important in daily life. Cataracts, diabetic retinopathy, conjunctivitis, glaucoma and other eye diseases are the most common eye diseases, they may even cause blindness in humans [1]. Cataract is the world's greatest cause of blindness. In our country there are 5 million cataract patients, nearly 1 million new patients each year. Prevention and treatment of cataracts is a challenge we have to face [1].

In our previous experiments, computer classifiers have been proposed to improve the accuracy and efficiency of cataract diagnosis [2-4]. It can be more quickly and accurately than professional doctor to identify large quantities of fundus images. However, with the increase number of samples and parameters, artificial selection of features has seriously restricted the classification of performance. So we hope that through the deep learning which can find the most distinguished features from thousands of parameters automatically. This is the first time to use the deep learning advantage capabilities on feature extraction. This solution is significant. Fig.1 shows the original cataract fundus images in different grading.

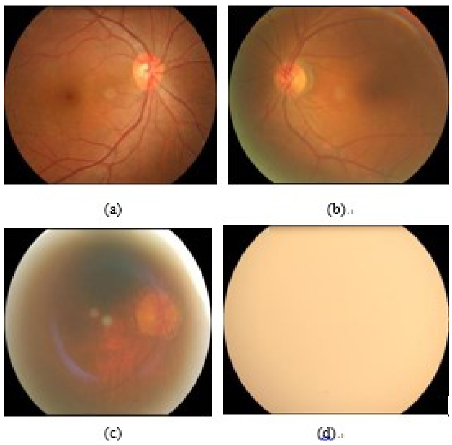


Fig.1four-class images of cataracts;

(a)normal;(b)mild;(c)medium;(d)severe

Only observe the degree of chaos on the fundus image, it is difficult to determine the degree of illness by the naked eye accurately.

In this paper, we still use the standard image classification process. Fig.2 shows the flow chart of complete classification identification process. It mainly contains three parts, which are fundus image preprocessing, feature extraction, and cataract classification and comparison[3-4]. Our research focuses on the feature extraction. As a comparison, we use retinal vascular feature and feature that is extracted from deep learning.

After feature extraction, SVM(support vector machine) and Softmax are used for cataract classification[4].The result in this paper proved that fundus image extracted feature through deep convolution network is effective and practical.

This paper is organized accordingly: Section2 describes related work; Section3 gives out the implementation, Section4 is experimental results and analysis. Section 5 is conclusion.

978-1-5386-1620-8/17/$31.00 ©2017 IEEE

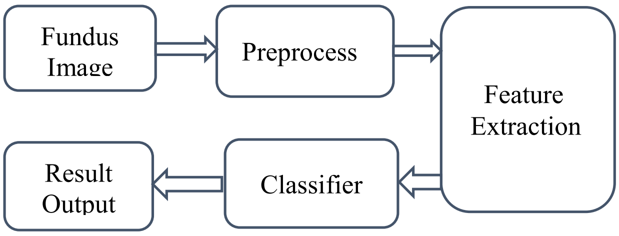


Fig.2 The flow chart of cataract classification.

II.RELATED WORK

The study of cataract has been going on for several years. Michael first applied the classifier to the cataract identify area. The retinal images were classified into normal and abnormal ones. Then the abnormal ones were given to the doctor [5]. Yang classified the retinal images into normal, mild, medium and severe ones [2]. Zeng had put forward ensemble learning algorithm (Bagging) and Random Forest to improve the accuracy of classification [4].But, he did not make the necessary preprocessing of images. Fan tried to use PCA (principal component analysis) to reduce the dimension of the image [9]. This method reduces the time of image classification, but it does not improve the accuracy of classification a lot. Dong used Kirsch's template filter and maximum entropy transformation to improve the accuracy of classification and greatly reduce the consume time of feature extraction [10]. The disadvantage of this method is that we do not know whether the features which we choose are suitable for our research.

We find that it is more and more difficult to use manually select features and general machine learning algorithm combination to obtain better result. Though classification of fundus image based on retinal vascular information has good performance last year[10]. At the same time, due to its good automatic feature selection ability and classification accuracy. deep learning is becoming more and more effective in different fields. So we decide to take action about deep learning and hope to get progress.

III IMPLEMENTATION

A. *Preprocessing*

Because of the different quality of our images. It is necessary for us to preprocessing the image. Image enhancement is a key step in the pre-process of fundus image. In the past, we used the histogram equalization to enhance quality of image, but it will lose a lot of image information, and don't highlight the blood vessels of the fundus image information. We select maximum entropy transformation after serval experiment On the basis of the principle of maximum entropy, the optimal classification threshold of gray level of the image is calculated by using the iterative algorithm, and then the local gray level transformation is performed by the transform function. Take the gray level of the images edge, by finding the best separation point, then get the threshold value on both sides and improve function of nonlinear stretching in the spatial domain(1-2). Through this method, not only the quality of the image has been enhanced, but also the information of the original fundus image is retained as much as possible. The formula is as follows

$$H(x)=\sum_{i=1}^{k} p(x=x_i)\log\frac{1}{p(x=x_i)} \qquad (1)$$

$$0 \le H(X) \le \log|X| \qquad (2)$$

B. *Deep Learning*

Deep learning was first introduced by Hinton[11] .It refers to the collection of algorithms that uses various machine learning algorithms on neural networks for various kinds of problems, Essentially, deep learning network is more complex CNN model. In this experiment, we use five convolution layers. After five convolution transformations, the image become higher level, more abstract representations. The deep learning model is built on Caffe which framework has excellent image performance.

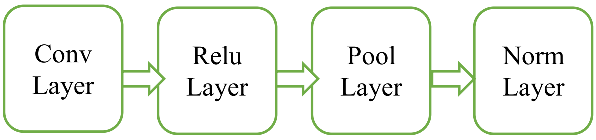


Fig.3 The model of convolution network layer

In Fig.3 above, this is a basic convolution network layer structure. Conv layer is feature extraction layer, we set first filter kernel size is 11, number output is 96, stride is 4, each neuron in the input and the previous layer of local receptive field are linked together, to extract the local characteristics. Once the local characteristics are extracted, the location of the relationship between it and other characteristics is identified. The feature of the convolution layer local connection and the weight sharing makes the calculation of the parameter much less (3-5).

Since there is no activation function on Conv layer, we choose to add the Relu layer as the activation function after Conv layer. The Relu function is used to add non-linear factors in order to provide non-linear modeling capabilities of the network, so that the network has a nonlinear mapping learning ability .At same time, it also solve the problem of gradient dissipation of BP algorithm in optimizing deep neural network (6). This step can help us fully extract non-linear features.

The pooling layer is mainly introduced by compressing the input feature map. On the one hand, the feature map becomes smaller, simplifies the computational complexity of the network. On the other hand, the feature compression is helpful to extract the main features. In this step, we select max pool, set the pool size is 3, stride is 2. Finally, we used the Norm layer to normalize the input region of the image. It can make the model have a better learning rate, as soon as reducing the possibility of overfitting. It make the extracted features more standardized. The above description uses formula as follow.

$$Y(m,n) = X(m,n) * H(m,n)$$
$$= \sum_{-\infty}^{+\infty}\sum_{-\infty}^{+\infty} X\binom{m-n}{i,n-j} H(i,j) \quad (3)$$

$$Y^l(m,n) = X^k(m,n) * H^{kl}(m,n)$$
$$= \sum_{k=0}^{K-1}\sum_{i=0}^{I-1}\sum_{j=0}^{J-1} X^k(m+i,n+j)H^{kl}(i,j) \quad (4)$$

$$\text{Calculations} = I * J * M * N * K * L \quad (5)$$
$$\varphi(X) = \max(0,x) \quad (6)$$

In the above formula, X (m, n) indicates normal image, H (m, n) represents convolutional kernel. Y (m, n) is expressed as a convolutional image. K.L indicate that the convolution layer has L output channels and K input channels respectively. In the formula (5), the size of the convolution kernel is I * J, and the size of the output channel is M * N. Formula (6) is the activation function in the neural network.

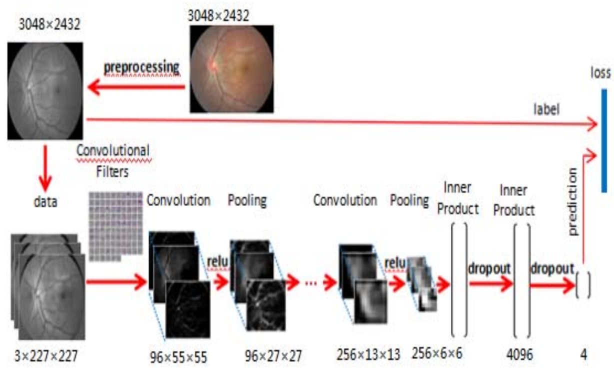
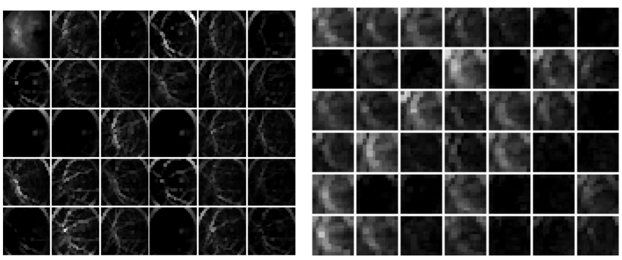


Fig.4 The model of deep learning

As shown above Fig. 4, we use five convolution layers as describe above. After five layers of convolved and pooling, the image was abstracted from 3048 * 2432 to 256 * 6 * 6 feature maps. As Fig.5 described below, fc3, fc4, fc5 represent the third, fourth, fifth of the deep learning network respectively.



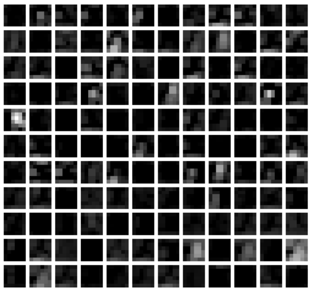


Fig.5 The image is extracted from fc3, fc4, fc5

As we all known, building a machine learning system, needs experienced experts to design feature extractor, and convert raw data into the appropriate feature vector. However deep learning don't need to extract the characteristics of the artificial design. This method has good generalization ability and robustness. Here we will derived all the data of fc5 layer, as the features of the classification tasks in the next stage. The data extracted from fc5 is huge matrix, and we do not know the exact meaning of the numbers in the matrix[12]. The following Fig.5 is the distribution of the feature map we extracted.

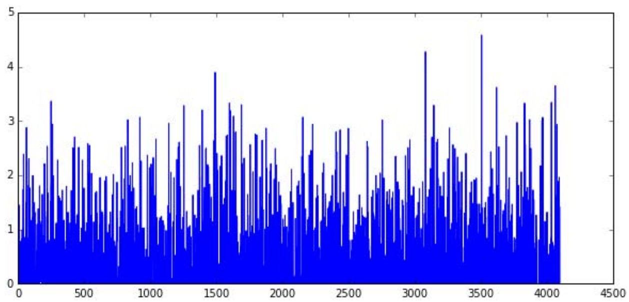


Fig.5 Feature map distribution

IV. EXPERIMENTAL RESULTS AND ANALYSIS

Retinal vascular information was extracted from fundus image which was initially experimented by Dong [10]. By this method, the wavelet feature classification accuracy we obtained was up to 84.7%. It has been proved wavelet feature is best feature before we extract features through deep learning. Here, we think it is appropriate to choose this feature as experimental comparison group.

After the step of feature extract, SVM, Softmax are used to classify the features that are extracted from fc5 of deep learning framework and wavelet feature from retinal vascular information. (Here, we set the feature extracted from deep learning as feature A, wavelet feature extracted from retinal vascular as feature B)

SVM is a very popular learning algorithm, which can be used for classification and regression, the principle is mapping the low dimensional input space into higher dimension one, then we use the kernel function which is determined through test set, to obtain the relaxation coefficient and the penalty coefficient. Last, we use test data to test accuracy of the classification.

In the below TABLE1-4, the accuracy is equal to the total amount of the correctly classified images in the test set divided by the total amount of the image to be classified. Normal, Slight, Medium, Severe are expressed as the corresponding accuracy. The symbols behind the numbers are the percent (%).

TABLE1.THE FOUR CLASSIFICATION RESULTS OF SVM CLASSIFIER

Features	Normal	Slight	Medium	Severe	accuracy
A	88.92	80.22	87.2	81.63	82.94
B	90.91	76.81	90.47	55.3	84.7

TABLE2.THE TWO CLASSIFICATION RESULTS OF SVM CLASSIFIER

Features	Normal	Non-Normal	accuracy
A	87.82	91.06	89.83
B	91.95	83.88	86

Softmax function is a generalization of the logistic function that compress K-dimensional vector of arbitrary values to a K-dimensional vector of values in the range [0, 1], the total sum is 1. The output of the Softmax function can be used to represent a categorical distribution. Here, the function calculates the accuracy of the two and four classes, respectively.

TABLE3.THE FOUR CLASSIFICATION RESULTS OF SOFTMAX CLASSIFIER

Features	Normal	Slight	Medium	Severe	accuracy
A	95.63	83.23	91.04	82.61	90.82
B	87.99	90.12	77.82	68.75	84.17

TABLE4.THE TWO CLASSIFICATION RESULTS OF SOFTMAX CLASSIFIER

Features	Normal	Non-Normal	accuracy
A	93.95	95.84	94.07
B	82.07	80.85	81.91

We collected 7851 fundus images, the normal is 4671, the mild is 2176, the medium is 622 and the severe is 382. These fundus images are classified by the professional doctor working in the department of ophthalmology, 70% randomly selected samples are used for training. The remaining 30% are used to test the accuracy of classifiers. After repeated training and testing for 50 times, we obtained the overall performance of the classification.

With comparing Table1 and Table3 , we know that the highest overall mean value is 90.82%, which are Softmax classifier results of feature A , 1% higher than the result classified by SVM, 6.1% , 6.6% higher than the correct rate of feature B classified by SVM and Softmax which value is 84.7% and 78.17%. At the same time, Tab.2 shows the feature A and feature B get the accuracy rates of 89.82% and 86% by SVM two-class classifier, and Tab.4 shows that the feature A and feature B get those of 94.01% and 81.91% respectively by Softmax two-class classifier. Through the above discussion we know the best accuracy rate of fundus four-classification is 90.82%. The best accuracy rate of two-classification is 94.07%.It has greatly improved in accuracy rate compared with previous work [7-11].

Meanwhile, the number of data set of our research reach 7851 fundus images. Around 20 times of our former work. In order to increase the risk of overfitting, we take 10 fold cross-validation, then calculate their average. This step is benefit improve the reliability of our experimental results.

However, because of the depth of convolution layer, even though we insert K80 GPU in server, the time spent on learning classification process, far more than the classification of vascular feature. Because the high quality of the image, training the existing sample has taken quite a long time, and if the number of images increases dramatically, this will cause us great difficulties.

Ⅴ. CONCLUSION

The deep learning network based on the five-layer convolution layer successfully separate the characteristics of the various nonlinear combinations in the fundus image by a multi-layer nonlinear mapping. Then extracting feature which is multilayer nonlinear mapping after processed from the last convolution layers .By comparing the results of the previous section of the experiment we found the feature that was extracted from deep learning more suitable for distinguishing the degree of cataract lesions in the training set. Though this feature is a huge matrix mathematically, and even we do not know what the characteristics of the matrix it represents. It has great improve the accuracy of classification. Despite we do not have enough large amount of images. We believe that with the advance of theory, release of more excellent open source framework, we can get the result of classification with quality of the upgrade.

In the future, we will continue our research on deep learning area. Furthermore, we will learn from Google's artificial intelligence framework to optimize our research.

REFERENCE

[1] Joes, Associate Member & IEEE Michael D.Abramoff. “Ridge based vessel segmentation in color images of the retina[J]”,IEEE Transactions on Medical Image,Vol.23, No.4,pp.501-509,Year 2004

[2] Meimei Yang, Jijiang Yang, Qinyan Zhang, Yu Niu, JianQiang Li, ”Classification of retinal image for automatic cataract detection”, e-Health Networking, Application & Service (Health com), 2013 IEEE 15th InternationalConference,Oct.9,pp.674-679,Year 2013.

[3] Liye Guo , Ji-Jiang Yang, Lihui Peng, Jianqiang Li, “A computer-aided healthcare system for cataract classification and grading based on fundus image analysis ”, Computer in Industry, Vol.69, pp. 72-78, May 2015

[4] Deng Xiuqin, Xiong Yong, PENG Hong. Effective adaptive weighted median filter algorithm. Computer Engineering and Applications, 2009，45(35)：185-187.

[5] Michael.Simons “Automated Detection of Anatomical Structures in Retinal Images”, International Conference Vol.3,pp.164-168, Year 2007.

[6] José Ignacio Orlandoa , b and Mariana del Fresnoac, "Review Preprocessing Feature Extraction Technical For Retial Blood Vessel Segmentation In fundus Image" José Ignacio Orlando, Sep 04, 2014

[7] Athinarayanan and Srinath ," Multi class cervical cancer classification in pap smear images using hybird texture feature and fuzzy logical based SVM" International Journal of Recent Scientific Research Vol. 7, Issue, 2, pp. 8831-8837, February, 2016

[8] Yang Zeng," Classification of Cataract Fundus Images Based on Ensemble Learning Algorithm" submitted publication

[9] Fan Weiming, "principal component analysis based cataract grading and classification" e-Health Networking, Application & Service (Health com), 2015 IEEE 17th International Conference on, Oct.9,pp.674-679,Year 2015.

[10] Dong Yanyan," Classification of Cataract Fundus Image Based on Retinal Vascular Information" Smart-Health ICSH2016,Dec 26,Year 2016

[11] Hinton "A Summary of Deep Learning Development", IEEE Science Image,Vol.3, No.4,pp.501-509,Year 2011

[12] Chen Zhen, "Feature Extraction Algorithm Based on Evolutionary Depth Learning", Computer Science Image,Vol.42 , No.11 ,Year 2015.11